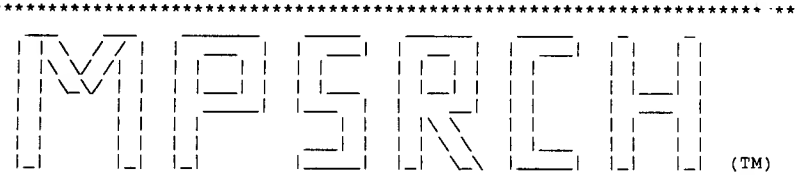

Release 3.1A John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat May 13 08:52:23 2000; MasPar time 3.67 Seconds
212.807 Million cell updates/sec
Tabular output not generated.

Title: >US-09-331-631-23
Description: (1-33) from US09331631.pep
Perfect Score: 287
Sequence: 1 RSGRGECRRQCLRRHEGQPWETQECMRRCRRRG 33

Scoring table: PAM 150
Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq35
1:geneseqp

Statistics: Mean 23.079; Variance 98.498; scale 0.234

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	287	100.0	33	1	W62836	Zea mays antimicrobial	3.17e-18
2	287	100.0	593	1	W62835	Zea mays antimicrobial	3.17e-18
3	269	93.7	35	1	R21079	Antimicrobial maize pe	1.63e-15
4	113	39.4	666	1	W62828	Macadamia integrifolia	2.95e-02
5	109	38.0	625	1	W62830	Macadamia integrifolia	6.47e-02
6	107	37.3	666	1	W62829	Macadamia integrifolia	9.56e-02
7	88	30.7	590	1	W62832	Gossypium hirsutum ant	3.61e+00
8	87	30.3	188	1	W00726	Vascular endothelial g	4.35e+00
9	87	30.3	188	1	W04829	Fibrosarcoma vascular	4.35e+00
10	87	30.3	188	1	W80493	Human vascular endothe	4.35e+00
11	86	30.0	525	1	W62831	Theobroma cacao antimi	5.25e+00
12	86	30.0	566	1	R20181	Sequence encoded by 67	5.25e+00
13	84	29.3	132	1	R28990	Encoded by clone TGR-C	7.60e+00
14	84	29.3	215	1	R28989	Encoded by GlcNAc rece	7.60e+00
15	80	27.9	179	1	W86221	Human VEGF-3 truncated	1.58e+01
16	80	27.9	184	1	W86220	Human VEGF-3 truncated	1.58e+01
17	80	27.9	189	1	W86219	Human VEGF-3 truncated	1.58e+01
18	80	27.9	194	1	W86218	Human VEGF-3 truncated	1.58e+01
19	80	27.9	201	1	W86236	Human VEGF-3 full leng	1.58e+01
20	80	27.9	221	1	W07611	Human vascular endothe	1.58e+01
21	80	27.9	221	1	W86205	Human vascular endothe	1.58e+01
22	80	27.9	388	1	Y04998	Mycobacterium species	1.58e+01
23	76	26.5	110	1	Y04944	Mycobacterium species	3.27e+01
24	74	25.8	55	1	W80489	Murine vascular endoth	4.68e+01
25	74	25.8	55	1	W04825	Vascular endothelial g	4.68e+01
26	74	25.8	145	1	W86213	Human VEGF-B truncated	4.68e+01
27	74	25.8	147	1	W86212	Human VEGF-B truncated	4.68e+01
28	74	25.8	150	1	W86211	Human VEGF-B truncated	4.68e+01
29	74	25.8	152	1	W86210	Human VEGF-B truncated	4.68e+01
30	74	25.8	155	1	W86209	Human VEGF-B truncated	4.68e+01
31	74	25.8	160	1	W86208	Human VEGF-B truncated	4.68e+01
32	74	25.8	167	1	W86234	Human VEGF-B full leng	4.68e+01
33	74	25.8	188	1	W00864	Murine VRF167.	4.68e+01
34	74	25.8	188	1	W04826	Heart vascular endothe	4.68e+01
35	74	25.8	188	1	W80490	Murine vascular endoth	4.68e+01
36	74	25.8	188	1	W86201	Human vascular endothe	4.68e+01
37	74	25.8	195	1	W04827	Heart vascular endothe	4.68e+01
38	74	25.8	195	1	W80491	Murine vascular endoth	4.68e+01
39	72	25.1	125	1	R28991	Encoded by clone TGR-C	6.68e+01
40	71	24.7	3025	1	P93284	Sequence of clone HIV-	7.98e+01
41	70	24.4	51	1	W33694	Mouse protamine 1.	9.51e+01
42	70	24.4	176	1	W33695	Mouse protamine 1 fuse	9.51e+01
43	70	24.4	387	1	W18663	Fragmented human NF-H	9.51e+01
44	70	24.4	554	1	W46873	Bacillus thuringiensis	9.51e+01
45	70	24.4	642	1	W05519	HCMV Toledo strain UL1	9.51e+01

ALIGNMENTS

RESULT 1
ID W62836 standard; Protein; 33 AA.
AC W62836;
DT 27-OCT-1998 (first entry)
DE Zea mays antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Zea mays.
PN WO9827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI; 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
PS Disclosure; Page 60; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 33 AA;

Query Match 100.0%; Score 287; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.17e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 RSGRGECRRQCLRRHEGQPWETQECMRRCRRRG 33
||||||||||||||||||||||||||||||||||
Qy 1 RSGRGECRRQCLRRHEGQPWETQECMRRCRRRG 33

RESULT 2
ID W62835 standard; Protein; 593 AA.
AC W62835;
DT 27-OCT-1998 (first entry)
DE Zea mays antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Zea mays.
PN WO9827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI; 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals

```
PS   Claim 1; Page 58-60; 96pp; English.
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
SQ   Sequence   593 AA;

  Query Match           100.0%;  Score 287;  DB 1;  Length 593;
  Best Local Similarity 100.0%;  Pred. No. 3.17e-18;
  Matches    33;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    561 RSGRGECRRQCLRRHEGQPWETQECMRRCRRRG 593
          |||||||||||||||||||||||||||||||||
Qy      1 RSGRGECRRQCLRRHEGQPWETQECMRRCRRRG 33


RESULT    3
ID   R21079 standard; Peptide; 35 AA.
AC   R21079;
DT   09-APR-1992  (first entry)
DE   Antimicrobial maize peptide, CMIII.
KW   Maize; CMIII; corn; pathogen.
OS   Zea mays.
PN   EP-465009-A.
PD   08-JAN-1992.
PF   05-JUN-1991; 305064.
PR   05-JUN-1990; US-536127.
PA   (PION-) PIONEER HI-BRED INT.
PI   Duvick JP, Rood TA, Rao AG;
DR   WPI; 92-010214/02.
PT   Use of maize seed peptide CMIII and DNA encoding it - for killing
PT   or inhibiting plant pathogenic microorganisms.
PS   Example 2; Page 5; 21pp; English.
CC   The peptide (SEQ ID NO 1) was purified from public corn variety B73
CC   and propriety corn variety MH18.  It is basic and has has a total
CC   mol. wt. of 3900 daltons.  The peptide sequence was used to design
CC   probes which were used to screen a maize genomic or cDNA library.
CC   The isolated CMIII gene can be used to prepare an expression vector
CC   for prodn. of recombinant CMIII for use in controlling plant patho-
CC   genic organisms.
CC   See also Q20272 and 3.
SQ   Sequence   35 AA;

  Query Match           93.7%;  Score 269;  DB 1;  Length 35;
  Best Local Similarity 97.0%;  Pred. No. 1.63e-16;
  Matches    32;  Conservative    0;  Mismatches   1;  Indels   0;  Gaps   0;

Db      1 RSGRGECRRQCLRRHEGQPYETQECMRRCRRRG 33
          |||||||||||||||||| ||||||||||||
Qy      1 RSGRGECRRQCLRRHEGQPWETQECMRRCRRRG 33


RESULT    4
ID   W62828 standard; Protein; 666 AA.
AC   W62828;
DT   27-OCT-1998  (first entry)
DE   Macadamia integrifolia antimicrobial protein.
KW   antimicrobial protein; infestation; control.
OS   Macadamia integrifolia.
FH   Key             Location/Qualifiers
FT   Peptide         1. .28
FT                   /note= "signal peptide"
FT   Protein         29. .666
FT                   /note= "mature protein"
PN   WO9827805-A1.
PD   02-JUL-1998.
PF   22-DEC-1997; AU0874.
PR   20-DEC-1996; AU-004275.
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI   Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR   WPI; 98-377279/32.
DR   N-PSDB; V42310.
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
PS   Claim 1; Page 34-36; 96pp; English.
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
SQ   Sequence   666 AA;

  Query Match           39.4%;  Score 113;  DB 1;  Length 666;
  Best Local Similarity 44.4%;  Pred. No. 2.95e-02;
  Matches    12;  Conservative    7;  Mismatches   7;  Indels   1;  Gaps   1;

Db    128 QCQKHCQRR-ETEPRHMQTCQQRCERR 153
          :|:::| || | :|:  | | :|| ||
Qy      6 ECRRQCLRRHEGQPWETQECMRRCRRR 32


RESULT    5
ID   W62830 standard; Protein; 625 AA.
AC   W62830;
DT   27-OCT-1998  (first entry)
DE   Macadamia integrifolia antimicrobial protein.
KW   antimicrobial protein; infestation; control.
OS   Macadamia integrifolia.
FH   Key             Location/Qualifiers
FT   Peptide         1. .28
FT                   /note= "signal peptide"
FT   Protein         29. .666
FT                   /note= "mature protein"
PN   WO9827805-A1.
PD   02-JUL-1998.
PF   22-DEC-1997; AU0874.
PR   20-DEC-1996; AU-004275.
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI   Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR   WPI; 98-377279/32.
DR   N-PSDB; V42316.
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
PS   Claim 1; Page 43-45; 96pp; English.
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
SQ   Sequence   625 AA;

  Query Match           38.0%;  Score 109;  DB 1;  Length 625;
  Best Local Similarity 44.0%;  Pred. No. 6.47e-02;
  Matches    11;  Conservative   10;  Mismatches   3;  Indels   1;  Gaps   1;

Db     43 QCQRRC-RQQESDPRQQQYCQRRCK 66
          :|:|:| .|:::|::|:: | | |||:
Qy      6 ECRRQCLRRHEGQPWETQECMRRCR 30


RESULT    6
ID   W62829 standard; Protein; 666 AA.
AC   W62829;
DT   27-OCT-1998  (first entry)
DE   Macadamia integrifolia antimicrobial protein.
KW   antimicrobial protein; infestation; control.
OS   Macadamia integrifolia.
FH   Key             Location/Qualifiers
FT   Peptide         1. .28
FT                   /note= "signal peptide"
FT   Protein         29. .666
FT                   /note= "mature protein"
PN   WO9827805-A1.
PD   02-JUL-1998.
PF   22-DEC-1997; AU0874.
PR   20-DEC-1996; AU-004275.
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI   Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR   WPI; 98-377279/32.
DR   N-PSDB; V42311.
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
```

```
PT   useful for controlling microbial infestations of plants or mammals
PS   Claim 1; Page 39-41; 96pp; English.
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
SQ   Sequence   666 AA;

  Query Match           37.3%;  Score 107;  DB 1;  Length 666;
  Best Local Similarity  44.4%;  Pred. No. 9.56e-02;
  Matches    12;  Conservative    7;  Mismatches   7;  Indels   1;  Gaps   1;

Db    128 QCQERC-QRHETEPRHMQTCQQRCERR 153
          :|: :| :||| :|:  | | :|| ||
Qy      6 ECRRQCLRRHEGQPWETQECMRRCRRR 32


RESULT    7
ID   W62832 standard; Protein; 590 AA.
AC   W62832;
DT   27-OCT-1998  (first entry)
DE   Gossypium hirsutum antimicrobial protein.
KW   antimicrobial protein; infestation; control.
OS   Gossypium hirsutum.
PN   WO9827805-A1.
PD   02-JUL-1998.
PF   22-DEC-1997; AU0874.
PR   20-DEC-1996; AU-004275.
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI   Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR   WPI; 98-377279/32.
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
PS   Claim 1; Page 49-51; 96pp; English.
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
SQ   Sequence   590 AA;

  Query Match           30.7%;  Score 88;  DB 1;  Length 590;
  Best Local Similarity  37.0%;  Pred. No. 3.61e+00;
  Matches    10;  Conservative   10;  Mismatches   6;  Indels   1;  Gaps   1;

Db    126 ECQQHC-HQQEQRPEKKQQCVRECREK 151
          ||:::| :::| :|   |:|:| || :
Qy      6 ECRRQCLRRHEGQPWETQECMRRCRRR 32


RESULT    8
ID   W00726 standard; Protein; 188 AA.
AC   W00726;
DT   30-NOV-1996  (first entry)
DE   Vascular endothelial growth factor-like protein SOM175-e6.
KW   Vascular endothelial growth factor; VEGF; SOM175-e6; neuron;
KW   astroglial proliferation.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   peptide         1. .21
FT                   /label= Sig_peptide
PN   WO9627007-A1.
PD   06-SEP-1996.
PF   22-FEB-1996; AU0094.
PR   02-MAR-1995; AU-001457.
PR   20-NOV-1995; AU-006647.
PR   22-DEC-1995; AU-007274.
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
PI   Grimmond S,  Hayward NK,  Larsson C,  Nordenskjold M;
PI   Weber G;
DR   WPI; 96-412774/41.
DR   N-PSDB; T33611.
PT   New growth factor related to vascular endothelial growth factor -
PT   useful for inducing astroglial proliferation and promoting neuronal
PT   survival
PS   Claim 12; Page 42-43; 113pp; English.
CC   Splice variants (W00726-28) of the human vascular endothelial growth
CC   factor-like polypeptide SOM175 (see also W00725) are products of
CC   cDNA clones (see also T33611-13) respectively lacking exon 6, exons
CC   6+7, and exon 4 of the SOM175 gene (see also T33610).  They show at
CC   least 1 of the properties of SOM175 including the ability to induce
CC   proliferation of vascular endothelial cells, to interact with
CC   flt-1/flk-1 receptors, and to induce cell migration, cell survival
CC   and/or an increase in intracellular levels of alkaline phosphatase.
CC   Recombinant SOM175 proteins can be used to induce astroglial
CC   proliferation and to promote neural survival and/or proliferation.
SQ   Sequence   188 AA;

  Query Match           30.3%;  Score 87;  DB 1;  Length 188;
  Best Local Similarity  40.6%;  Pred. No. 4.35e+00;
  Matches    13;  Conservative    7;  Mismatches   9;  Indels   3;  Gaps   3;

Db    137 SPRPLCPR-CTQHHQ-RP-DPRTCRCRCRRRS 165
          | |  | | | ::|: :| : : |  |||||:
Qy      2 SGRGECRRQCLRRHEGQPWETQECMRRCRRRG 33


RESULT    9
ID   W04829 standard; Protein; 188 AA.
AC   W04829;
DT   28-APR-1997  (first entry)
DE   Fibrosarcoma vascular endothelial growth factor-B167.
KW   Endothelial cell; proliferation; vascular endothelial growth factor; VPF;
KW   VEGF; endothelium; mesodermal cell; cationic dimer; tissue regeneration;
KW   vascular permeability factor; cell mitogen; angiogenesis; cell growth;
KW   embryonic development; wound healing; tissue reorganisation; antibody;
KW   cancer; metastatic risk; tumour cell; mouse.
OS   Homo sapiens.
PN   WO9626736-A1.
PD   06-SEP-1996.
PF   01-MAR-1996; U02957.
PR   01-MAR-1995; US-397651.
PR   06-JUN-1995; US-469427.
PR   06-DEC-1995; US-569063.
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (UYHE-) UNIV HELSINKI LICENSING LTD OY.
PI   Alitalo K,  Eriksson U,  Olofsson B,  Pajusola K;
DR   WPI; 96-412582/41.
DR   N-PSDB; T37913.
PT   Vascular endothelial growth factor VEGF-B proteins - useful to
PT   accelerate angiogenesis in wound healing, also related nucleic acid
PT   and antibodies for cancer diagnosis
PS   Claim 18; Page 59; 107pp; English.
CC   W04824-W04831 represent the vascular endothelial growth factor (VEGF)
CC   proteins of the invention, which promote endothelial or mesodermal cell
CC   proliferation.  VEGF is also a glycosylated cationic dimer, and is
CC   sometimes referred to as vascular permeability factor (VPF).  VEGF has
CC   diverse effects, depending on the specific biological context in which it
CC   is found.  VEGF is a potent endothelial cell mitogen, and directly
CC   contributes to induction of angiogenesis in vivo by promoting endothelial
CC   cell growth during normal embryonic development, wound healing, and
CC   tissue regeneration/reorganisation.  The VEGF proteins of the invention
CC   share the angiogenic and other properties of VEGF, but are distributed
CC   and expressed in tissues differently to VEGF.  The proteins can therefore
CC   be used to accelerate angiogenesis in wound healing.  Antibodies against
CC   the proteins can be used for inhibiting angiogenesis.  The antibodies can
CC   also be used diagnostically to quantitatively detect VEGF-B.  Primers
CC   complementary to the coding sequences for the proteins of the invention
CC   can also be used to detect VEGF-B coding sequences.  Quantification of
CC   VEGF-B in cancer biopsy specimens may be useful as an indicator of
CC   metastatic risk.  VEGF-B expression in a cell can be retarded using
CC   antisense sequences direct against the VEGF coding sequences, this is
CC   especially useful in retarding VEGF expression in tumour cells.
SQ   Sequence   188 AA;

  Query Match           30.3%;  Score 87;  DB 1;  Length 188;
  Best Local Similarity  40.6%;  Pred. No. 4.35e+00;
  Matches    13;  Conservative    7;  Mismatches   9;  Indels   3;  Gaps   3;
```

```
Db    137 SPRPLCPR-CTQHHQ-RP-DPRTCRCRCRRRS 165
          | |  | | | ::|: :| : : |  |||||:
Qy      2 SGRGECRRQCLRRHEGQPWETQECMRRCRRRG 33
```

RESULT 10
ID W80493 standard; Protein; 188 AA.
AC W80493;
DT 29-JAN-1999 (first entry)
DE Human vascular endothelial growth factor (VEGF)-B167.
KW Vascular endothelial growth factor; VEGF; proliferation; human;
KW endothelial cell; angiogenesis; tissue growth; organ repair.
OS Homo sapiens.
PN US5840693-A.
PD 24-NOV-1998.
PF 01-MAR-1996; 609443.
PR 01-MAR-1996; US-609443.
PR 01-MAR-1995; US-397651.
PR 06-JUN-1995; US-469427.
PR 06-DEC-1995; US-569063.
PA (LUDW-) LUDWIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.
PI Alitalo K, Eriksson U, Olofsson B, Pajusola K;
DR WPI; 99-034079/03.
DR N-PSDB; V63567.
PT Vascular endothelial growth factor-B iso:forms, and DNA encoding
PT them - useful for inducing angiogenesis and cellular proliferation,
PT and raising antibodies to inhibit activities in e.g. tumours
PS Claim 1; Fig 11; 52pp; English.
CC The present sequence represents human vascular endothelial growth factor
CC (VEGF)-B167. VEGF proteins are used for promoting proliferation of
CC endothelial cells and for stimulating angiogenesis (the proliferation
CC of new capillaries form pre-existing blood vessels). These activities
CC are useful for treating tissue growth and repair, including organ
CC repair. This is also useful in pregnancy, in follicle development,
CC as these processes must occur in development of the placenta. The
CC proteins can also be used to raise antibodies, either for use in
CC detection of the proteins or as inhibitors of their action. This is
CC especially useful as angiogenesis is required by tumours as they need
CC new blood supplies to grow and proliferate.
SQ Sequence 188 AA;

```
  Query Match           30.3%;  Score 87;  DB 1;  Length 188;
  Best Local Similarity 40.6%;  Pred. No. 4.35e+00;
  Matches   13;  Conservative    7;  Mismatches   9;  Indels   3;  Gaps   3;

Db    137 SPRPLCPR-CTQHHQ-RP-DPRTCRCRCRRRS 165
          | |  | | | ::|: :| : : |  |||||:
Qy      2 SGRGECRRQCLRRHEGQPWETQECMRRCRRRG 33
```

RESULT 11
ID W62831 standard; Protein; 525 AA.
AC W62831;
DT 27-OCT-1998 (first entry)
DE Theobroma cacao antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Theobroma cacao.
PN WO9827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI; 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 47-49; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 525 AA;

```
  Query Match           30.0%;  Score 86;  DB 1;  Length 525;
  Best Local Similarity 36.0%;  Pred. No. 5.25e+00;
  Matches    9;  Conservative   10;  Mismatches   5;  Indels   1;  Gaps   1;

Db     87 QCQGRCQEQQQGQ-REQQQCQRKCW 110
          :|: :|  :::|| :| |:| |:|:
Qy      6 ECRRQCLRRHEGQPWETQECMRRCR 30
```

RESULT 12
ID R20181 standard; Protein; 566 AA.
AC R20181;
DT 16-APR-1992 (first entry)
DE Sequence encoded by 67 kD T. cacao protein cDNA.
KW Cocoa; flavour; vicilin; seed storage protein.
OS Theobroma cacao.
PN WO9119801-A.
PD 26-DEC-1991.
PF 07-JUN-1991; G00914.
PR 11-JUN-1990; GB-013016.
PA (MRSC) MARS UK LTD.
PI Spencer ME, Hodge R, Deakin EA, Ashton S;
DR WPI; 92-024418/03.
DR N-PSDB; Q20377.
PT Recombinant cocoa proteins - are responsible for flavour in cocoa
PT beans and produced in large quantities using yeast and bacterial
PT expression vectors
PS Claim 4; Fig 2; 59pp; English.
CC The inventors claim a 67 kD and 31 kD T. cacao protein, and
CC fragments, and encoding DNAs. The 47 kD and 31 kD proteins are
CC derived from the 67 kD precursor. T. cacao protein cDNA was
CC detected in a cDNA library prepared from immature cocoa beans RNA
CC using a probe based on the AA sequence of a CNBr peptide common to
CC the 47 kD and 31 kD polypeptides. Homology searches revealed close
CC homologies between the 67 kD polypeptide and the vicilins, which are
CC seed storage proteins.
SQ Sequence 566 AA;

```
  Query Match           30.0%;  Score 86;  DB 1;  Length 566;
  Best Local Similarity 36.0%;  Pred. No. 5.25e+00;
  Matches    9;  Conservative   10;  Mismatches   5;  Indels   1;  Gaps   1;

Db     87 QCQGRCQEQQQGQ-REQQQCQRKCW 110
          :|: :|  :::|| :| |:| |:|:
Qy      6 ECRRQCLRRHEGQPWETQECMRRCR 30
```

RESULT 13
ID R28990 standard; Protein; 132 AA.
AC R28990;
DT 06-APR-1993 (first entry)
DE Encoded by clone TGR-CL1 complementary to TGR-CL1bis.
KW Thyroid N-acetyl-glucosamine receptor; carbohydrate recognition domain;
KW GlcNAc receptor; CRD; thyroid adenocarcinoma; lysosome; endosome;
KW Hashimoto's disease; Basedow's disease.
OS Homo sapiens.
PN WO9219733-A.
PD 12-NOV-1992.
PF 30-APR-1992; F00396.
PR 03-MAY-1991; FR-005478.
PA (CNRS) CENT NAT RECH SCI.
PI Blanck O, Courageot J, Miquelis R, Thibault V;
DR WPI; 92-398860/48.
DR N-PSDB; Q31014.
PT New thyroid N-acetyl-glucosamine receptor proteins - and
PT fragments, antibodies and encoding nucleic acids, for therapeutic
PT or diagnostic use
PS Disclosure; Fig 5a; 75pp; French.
CC Clone TGR-CL1bis (see Q31005) was isolated by immunoscreening a
CC Western blot derived from a cDNA bank of normal human thyroid cDNA
CC in lambda gt11. Denatured GlcNAc receptor was purified from porcine
CC thyroids and used to raise polyclonal antibodies in rabbits for use
CC in immunoscreening. Sequence TGR-CL1 is complementary to

```
CC   the corresponding "bis" sequence and includes an ORF. The amino acid
CC   sequence deduced from the complementary sequence is not a GlcNAc
CC   receptor. See Q31005-Q31018.
SQ   Sequence   132 AA;

  Query Match           29.3%;  Score 84;  DB 1;  Length 132;
  Best Local Similarity  36.7%;  Pred. No. 7.60e+00;
  Matches    11;  Conservative    6;  Mismatches  10;  Indels   3;  Gaps   3;

Db     27 GRSPCRHPCAPRQSQCAPWSGP-CAQ-CWR 54
          ||: ||: |  |: :  ||    | : |:|
Qy      3 GRGECRRQCLRRH-EGQPWETQECMRRCRR 31


RESULT   14
ID   R28989 standard; Protein; 215 AA.
AC   R28989;
DT   06-APR-1993  (first entry)
DE   Encoded by GlcNAc receptor complementary sequence.
KW   Thyroid N-acetyl-glucosamine receptor; carbohydrate recognition domain;
KW   CRD; thyroid adenocarcinoma; lysosome; endosome; TGR-CL1; TGR-CL5;
KW   Hashimoto's disease; Basedow's disease.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   misc_difference 50
FT                   /note= "corresponds to codon TGS"
FT   region          82. .127
FT                   /note= "overlap"
PN   WO9219733-A.
PD   12-NOV-1992.
PF   30-APR-1992; F00396.
PR   03-MAY-1991; FR-005478.
PA   (CNRS ) CENT NAT RECH SCI.
PI   Blanck O, Courageot J, Miquelis R, Thibault V;
DR   WPI; 92-398860/48.
DR   N-PSDB; R28989.
PT   New thyroid N-acetyl-glucosamine receptor proteins - and
PT   fragments, antibodies and encoding nucleic acids, for therapeutic
PT   or diagnostic use
PS   Disclosure; Fig 3; 75pp; French.
CC   Clones TGR-CL1bis and TGR-CL5bis (see Q31005 and Q31007,
CC   respectively) were isolated by immunoscreening a Western blot derived
CC   from a cDNA bank of normal human thyroid cDNA in lambda gt11.
CC   Denatured GlcNAc receptor was purified from porcine thyroids and
CC   used to raise polyclonal antibodies in rabbits for use in
CC   immunoscreening. Sequences TGR-CL1 and TGR-CL5 are complementary to
CC   the corresponding "bis" sequences and partially overlap one another.
CC   The amino acid sequence deduced from the combined overlapping
CC   sequences is not a GlcNAc receptor.
CC   See Q31005-Q31018.
SQ   Sequence   215 AA;

  Query Match           29.3%;  Score 84;  DB 1;  Length 215;
  Best Local Similarity  36.7%;  Pred. No. 7.60e+00;
  Matches    11;  Conservative    6;  Mismatches  10;  Indels   3;  Gaps   3;

Db    108 GRSPCRHPCAPRQSQCAPWSGP-CAQ-CWR 135
          ||: ||: |  |: :  ||    | : |:|
Qy      3 GRGECRRQCLRRH-EGQPWETQECMRRCRR 31


RESULT   15
ID   W86221 standard; protein; 179 AA.
AC   W86221;
DT   16-FEB-1999  (first entry)
DE   Human VEGF-3 truncated fragment 4.
KW   VEGF; VRP; vascular endothelial growth factor; VEGF-related protein;
KW   recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW   coronary; collateral vessel development; cell growth; migration; heart;
KW   lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW   wound healing; skin; vascular permeability.
OS   Homo sapiens.
PN   WO9849300-A2.
PD   05-NOV-1998.
PF   20-APR-1998; U07801.
PR   25-APR-1997; US-842984.
PA   (COLL-) COLLATERAL THERAPEUTICS.
PI   Bohlen P;
DR   WPI; 99-009426/01.
PT   New truncated vascular endothelial growth factor-related protein
PT   subunits - lack part of the N-terminal sequence, used to stimulate
PT   angiogenesis, e.g. for treating heart disease and ischaemia
PS   Claim 5; Fig 2C; 113pp; English.
CC   The invention relates to truncated VRP (vascular endothelial growth
CC   factor (VEGF)-related protein) subunits that have at least one amino acid
CC   N-terminal to the first Cys of the core sequence deleted. Host cells
CC   transformed or transfected with expression vectors containing nucleic
CC   acids encoding the truncated VRP subunits are used to produce the
CC   truncated proteins recombinantly. The truncated VRP subunits, optionally
CC   expressed from gene therapy vectors, have in vivo and in vitro angiogenic
CC   activity and are used to stimulate angiogenesis, particularly coronary
CC   collateral vessel development in cases of cardiac ischaemia; to stimulate
CC   endothelial cell growth and migration in vitro; to treat heart disease;
CC   to treat ischaemia (e.g.cardiac, chronic coronary or chronic lower limb
CC   ischaemia; stroke and peripheral vascular disease); to promote healing of
CC   wounds (of skin or intestines), and to increase vascular permeability.
CC   Sequences W86218 to W86221 represent truncated fragments of VEGF-3.
SQ   Sequence   179 AA;

  Query Match           27.9%;  Score 80;  DB 1;  Length 179;
  Best Local Similarity  40.6%;  Pred. No. 1.58e+01;
  Matches    13;  Conservative    6;  Mismatches  10;  Indels   3;  Gaps   3;

Db    128 SPRPLCPR-CTQHHQC-P-DPRTCRCRCRRRS 156
          | |  | | | ::|:  | : : |  |||||:
Qy      2 SGRGECRRQCLRRHEGQPWETQECMRRCRRRG 33


Search completed: Sat May 13 08:52:30 2000
Job time : 7 secs.
```